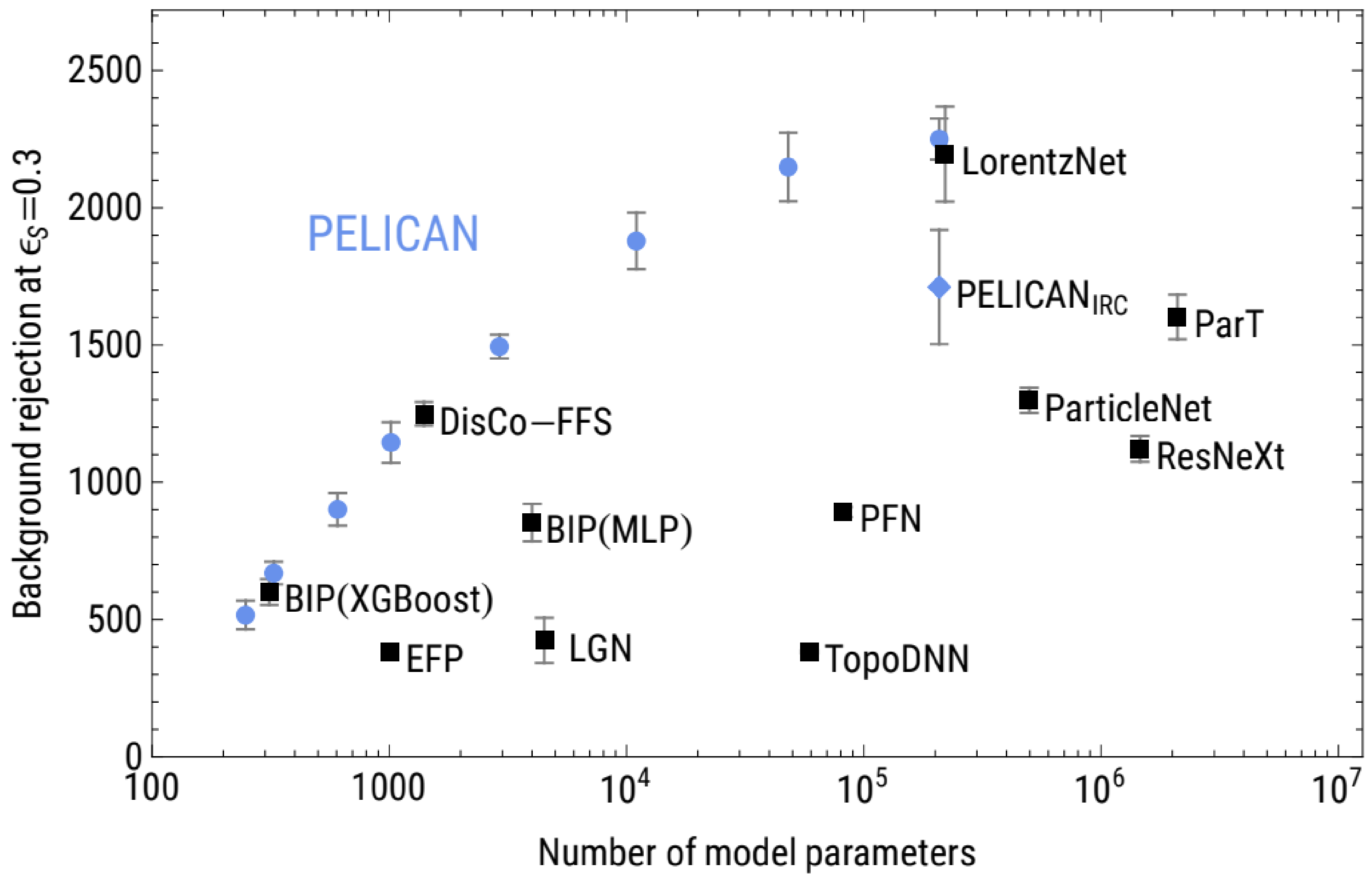
Classifying u/d jets using p_T weighted jet charge

Katherine Fraser, **Rabia Husain**, Noah McNeal, and Rashmish K. Mishra

HARVARD UNIVERSITY

Using ML to classify jets



- As jet tagging has improved, we have been able to reach higher background rejections with more parameters

Image from Bogatskiy et al., 2307.16506

Why is u/d especially hard?

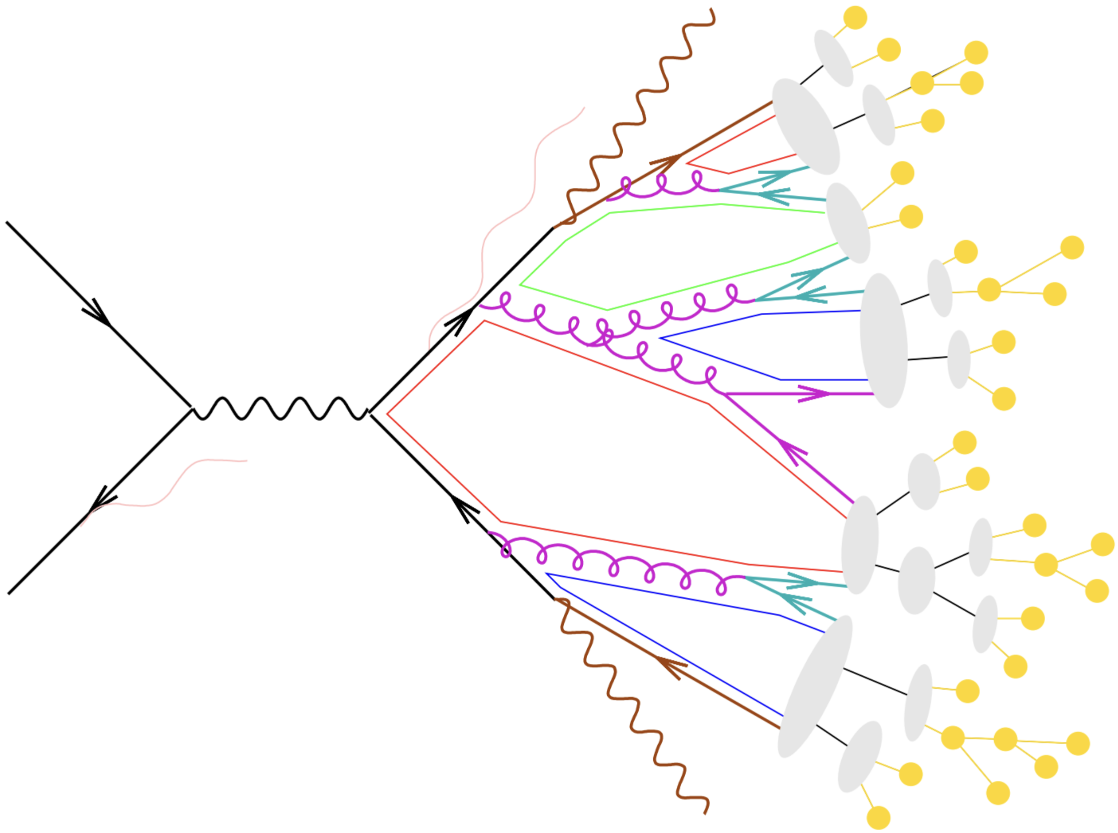


Image from D. Zeppenfeld, “Event generation and parton shower”, PiTP lecture (2005)

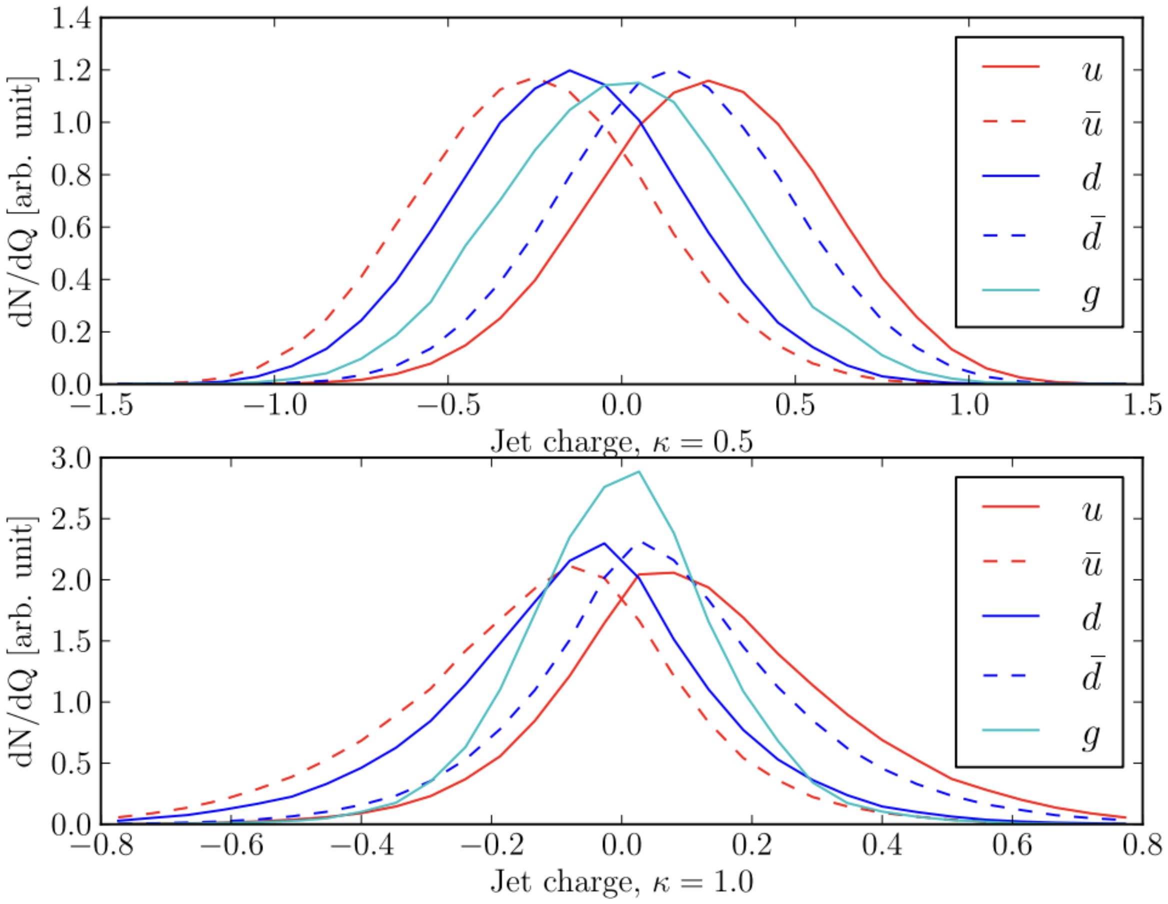
- Particles that initiate jets radiate and cascade, obscuring the features of the initiating particle
- u and d have the same SU(3) charge, so their jets look similar
- We can use electric charge to try to distinguish their jets

p_T weighted jet charge can help

Krohn et al., 1209.2421
Field and Feynman, 1977

$$\mathcal{Q}_\kappa^i = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa$$

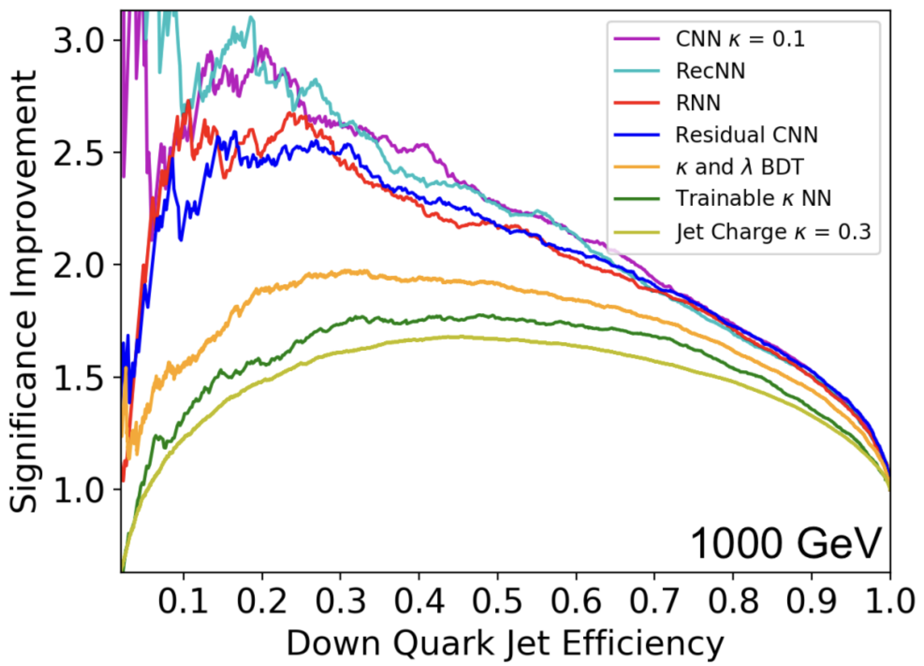
- Low κ enhances soft contributions and helps separate distributions

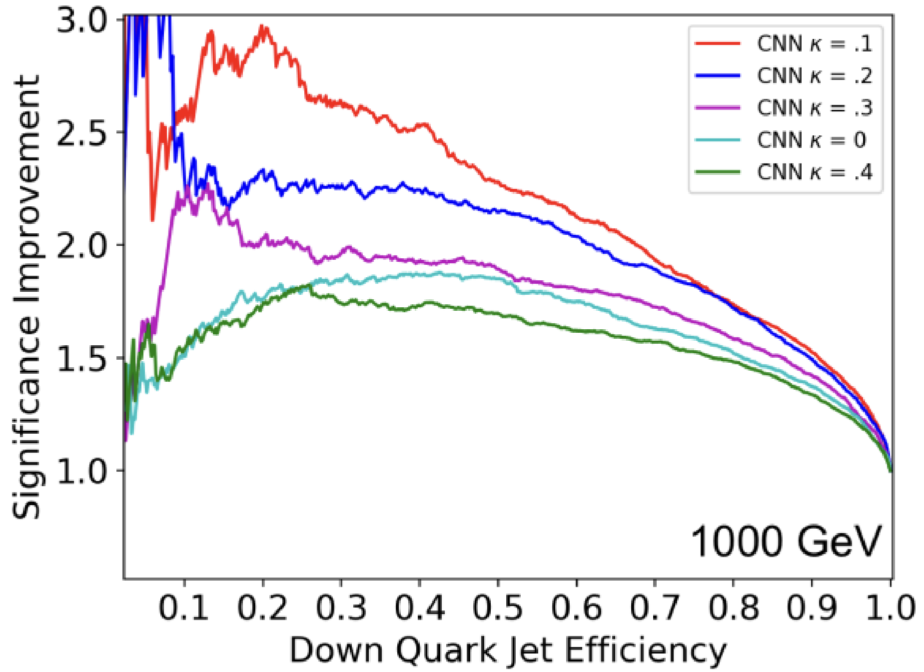


Older results

Fraser and Schwartz, 1803.08066

- Jet charge improves performance, but performance depends on κ
- $\mathrm{SI} = \epsilon_s/\sqrt{\epsilon_b}, \quad \epsilon_s = TPR, \quad \epsilon_b = FPR$





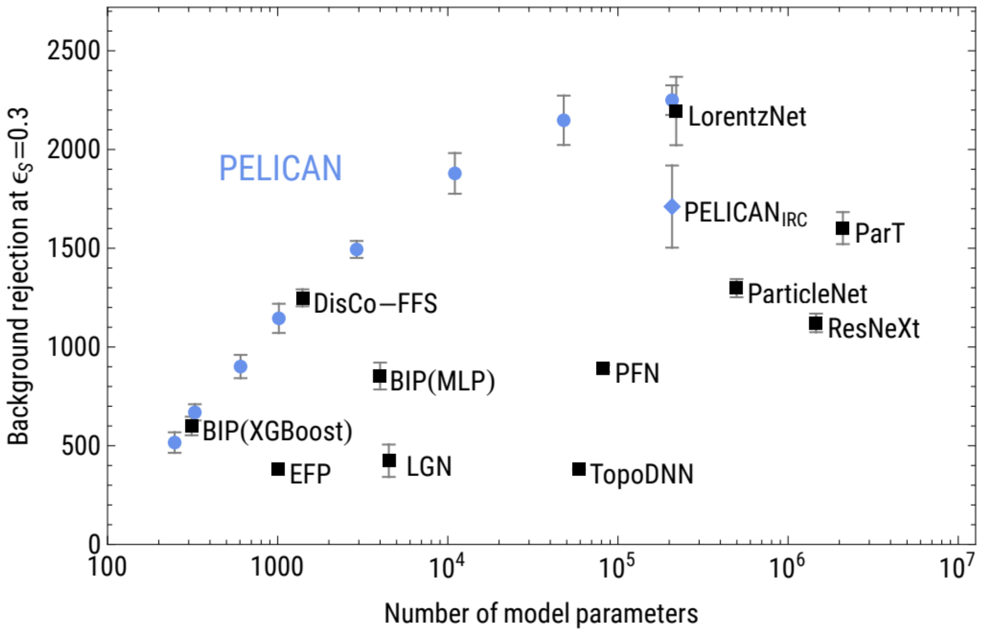
Newer Architectures

Bogatskiy et al., 2307.116506
Gong et al., 2201.08187
Qu et al., 2202.03772
Wu et al., 2407.08682

- Newer networks use more sophisticated techniques to improve jet tagging
- So far have not been applied to the u/d problem
- We choose four newer architectures:
  - PELICAN (GNN)
  - LorentzNet (GNN)
  - ParT (Transformer)
  - MI-ParT (Transformer)

- We give momenta and scalars as inputs to these networks

Image from Bogatskiy et al., 2307.16506



Methods

- Used sample of 2M jets generated using Pythia 8.311
- 1M up quark initiated and 1M down quark initiated
- Supervised training using labels from Pythia
- All training includes 4-momenta
  - LorentzNet and PELICAN take the pairwise dot products of the input 4-momenta and add two auxiliary beam particles
  - ParT and MI-ParT use kinematic and trajectory displacement features between particles, but we disregard these in our studies
- We tested many configurations of scalars that hold both particle and jet level information

Particle Level

- PID
- Charge
- Particle p_T weighted jet charge

$$Q_\kappa = \frac{1}{(p_T^{\text{jet}})^\kappa} Q (p_T)^\kappa$$

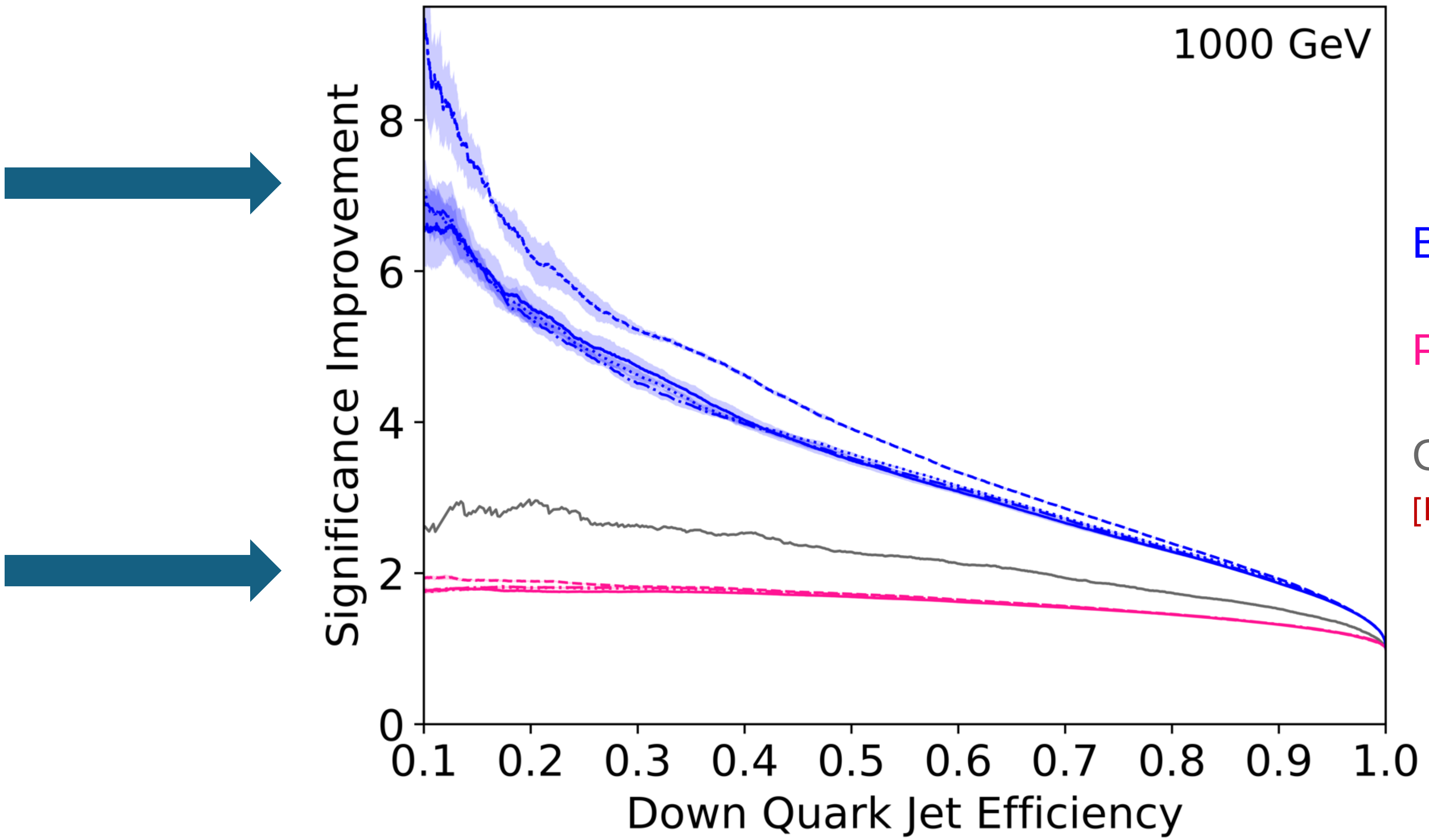
Jet Level

- Overall p_T weighted jet charge

$$Q_\kappa^i = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa$$

Takeaways

- Particle level information makes a difference if you have the right architecture



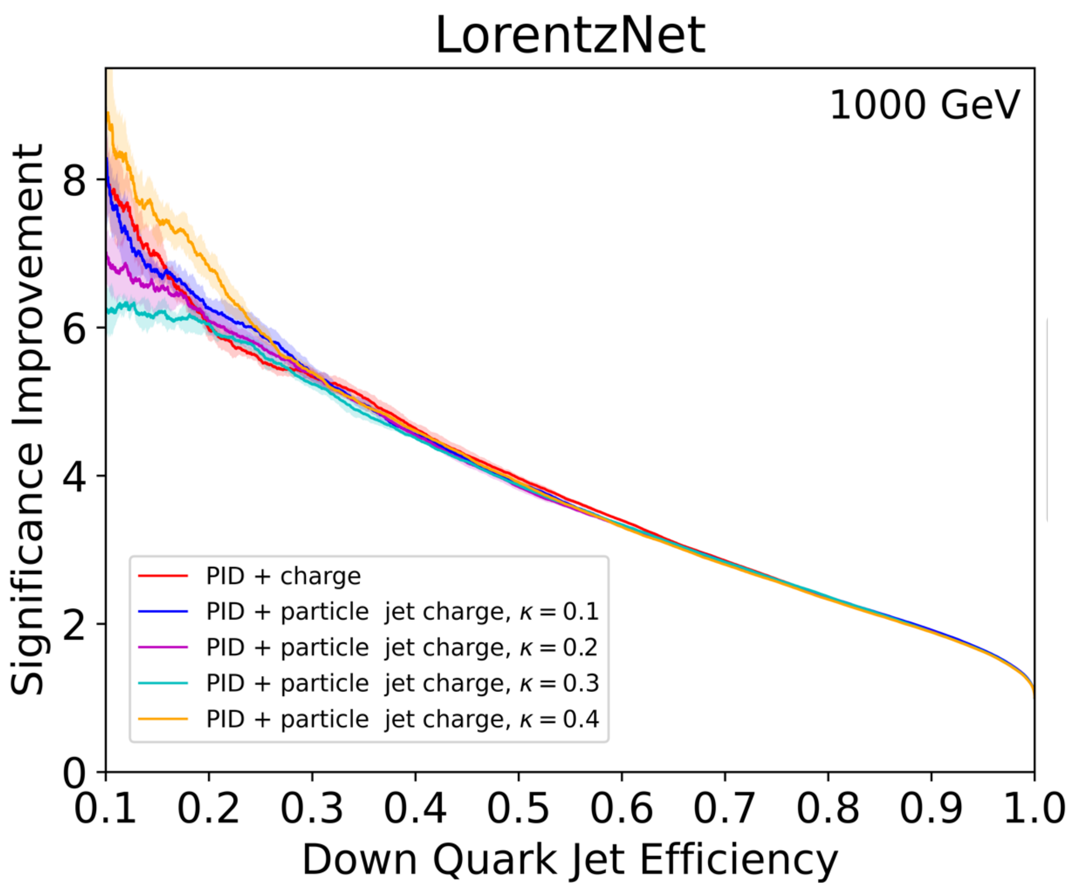
Blue – PID + overall jet charge, κ = 0.2

Pink – overall jet charge, κ = 0.2

Grey – CNN, κ = 0.1

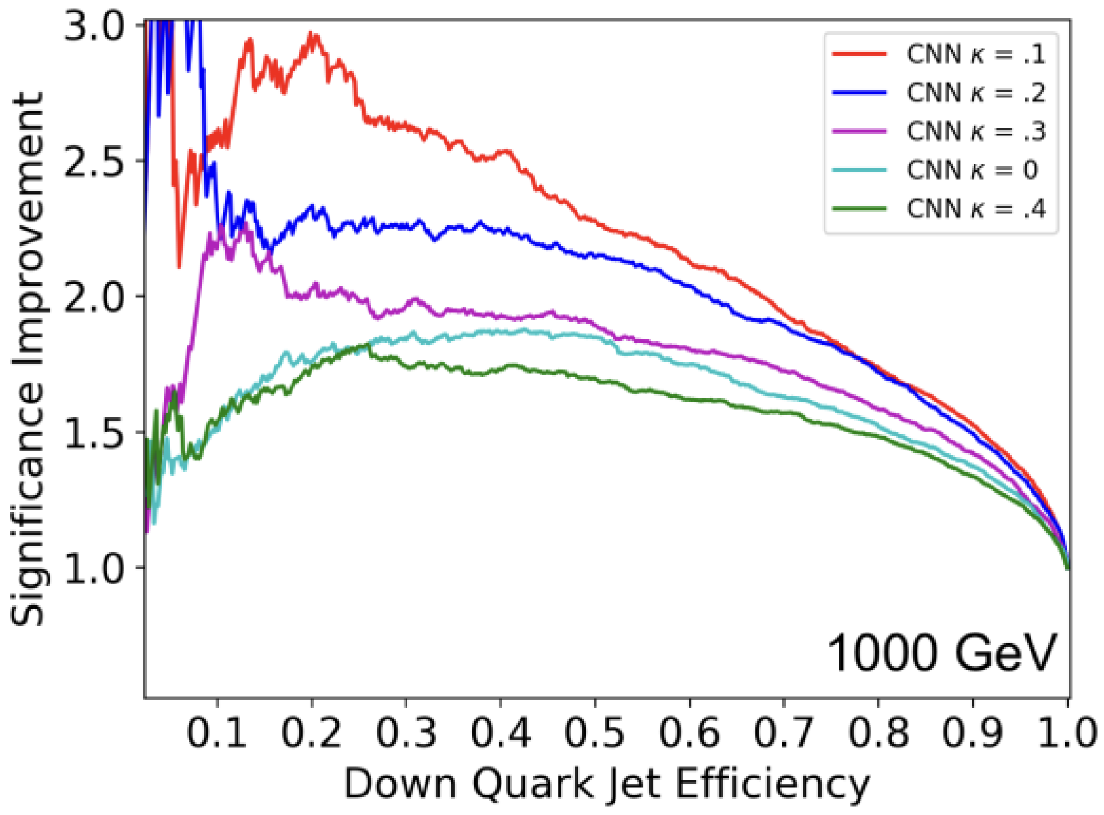
[Fraser and Schwartz, 1803.08066]

- Newer networks are independent of κ in particle jet charge unlike older networks
- This holds for all networks



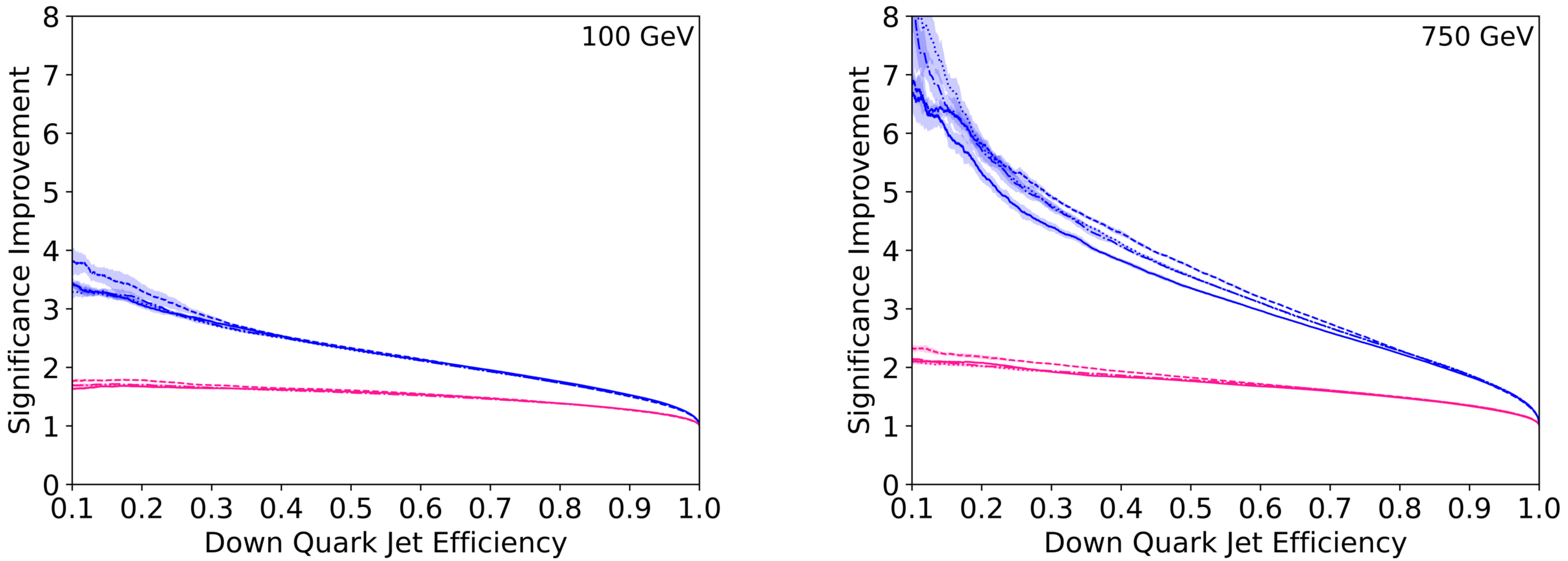
Our result

Fraser and Schwartz, 1803.08066



Previous result

Performance dependence on jet p_T



Blue – PID + charge overall jet charge, $\kappa = 0.3$

Pink – overall jet charge, $\kappa = 0.3$

Summary

- We see significant improvements to u/d tagging with the inclusion of particle level information in newer networks
- Results are no longer sensitive to the value of κ, the p_T weight
- Results hold when changing jet p_T, dataset size, network size
  - These affect the amount of separation between curves

Network	1000 GeV AUC
CNN*	0.879
PELICAN	0.923
LorentzNet	0.929
ParT	0.927
MI-ParT	0.925

Our values are quoted for:

PID + overall jet charge, $\kappa = 0.2$

*Fraser and Schwartz, 1803.08066

Backup

Older results

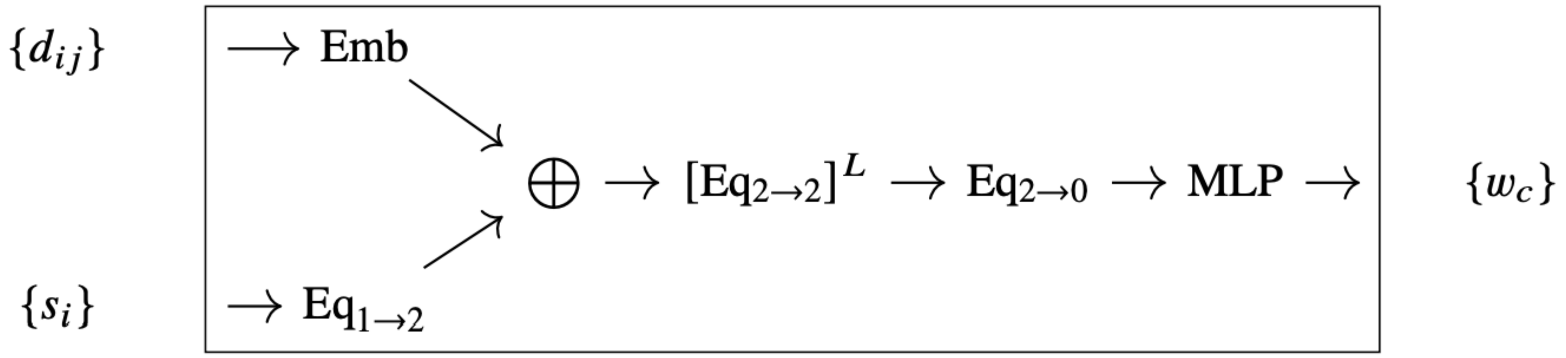
Fraser and Schwartz, 1803.08066

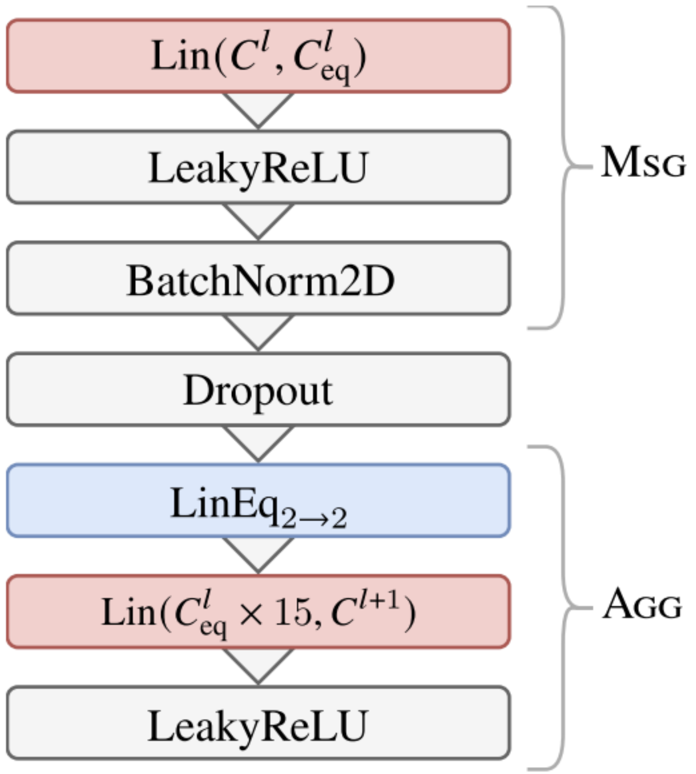
Network	100 GeV Up Quark Efficiency	100 GeV AUC	1000 GeV Up Quark Efficiency	1000 GeV AUC
RecNN	0.085	0.834	0.049	0.876
CNN	0.080	0.837	0.048	0.879
RNN	0.079	0.841	0.054	0.874
Residual CNN	0.078	0.840	0.053	0.877
κ and λ BDT	0.090	0.830	0.068	0.859
Trainable κ NN	0.104	0.815	0.080	0.841
Jet Charge	0.109	0.810	0.090	0.832

PELICAN

Bogatskiy et al., 2307.16506

- Permutation equivariant
- Complete set of Lorentz invariants at the input stage
  - Pairwise dot products of the input 4-momenta
  - Two auxiliary beam particles
  - Scalar data (PID, charge, etc)

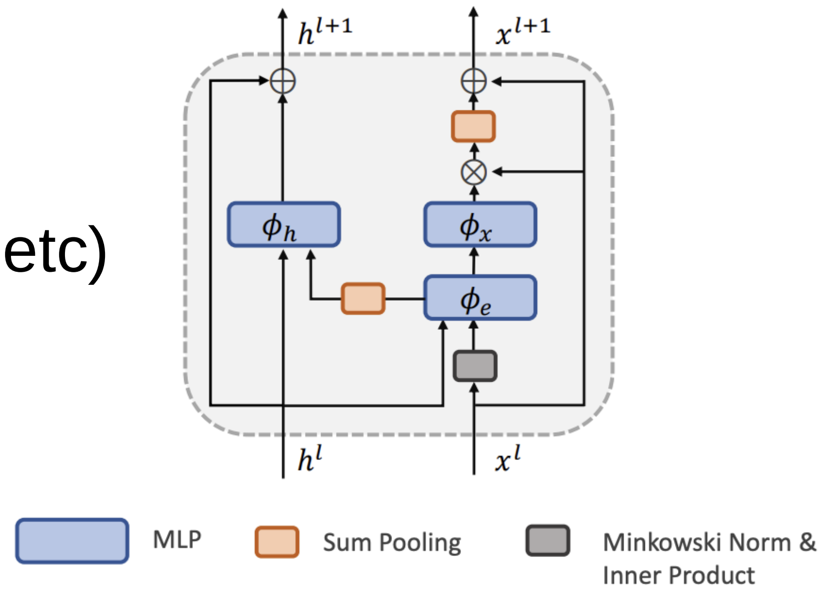




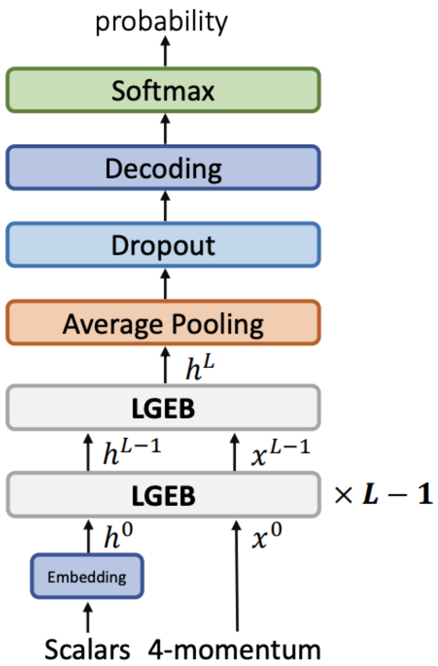
LorentzNet

Gong et al., 2201.08187

- Lorentz equivarant
  - Constructed using Minkowski dot products of the input 4-momenta
- Inputs:
  - Particle 4-momenta
  - Scalar data (PID, charge, etc)



Lorentz Group Equivariant Block (LGEB)

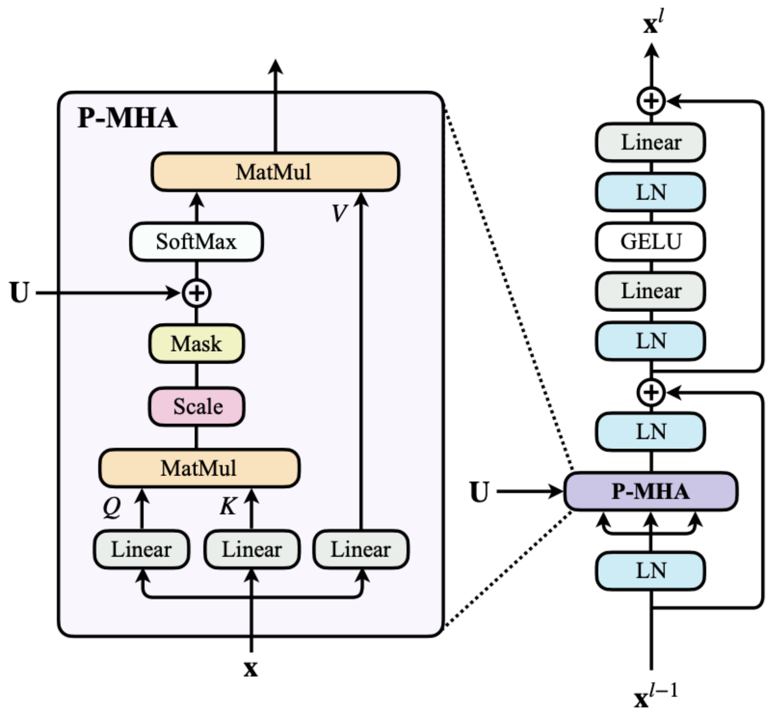


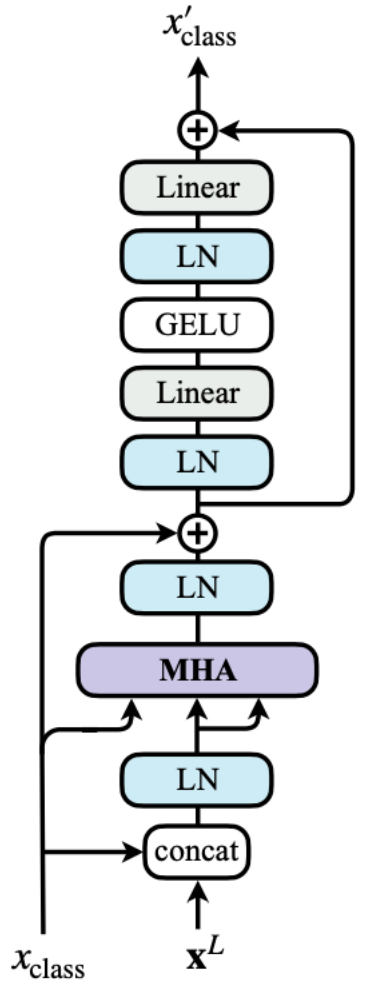
LorentzNet

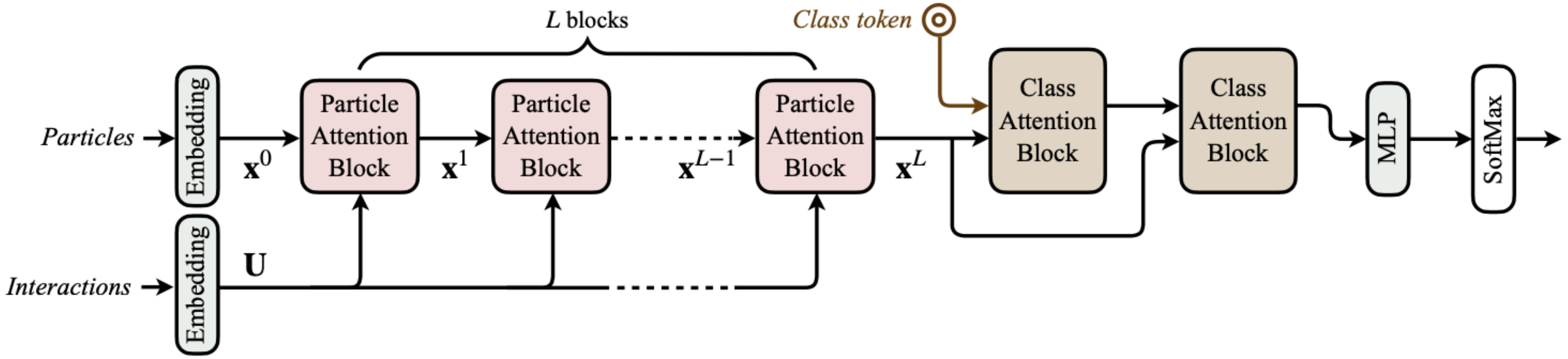
ParT

Qu et al., 2202.03772

- Transformer based architecture
- Incorporates pairwise particle interactions in attention
- Inputs:
    - Particle features
    - Interaction features



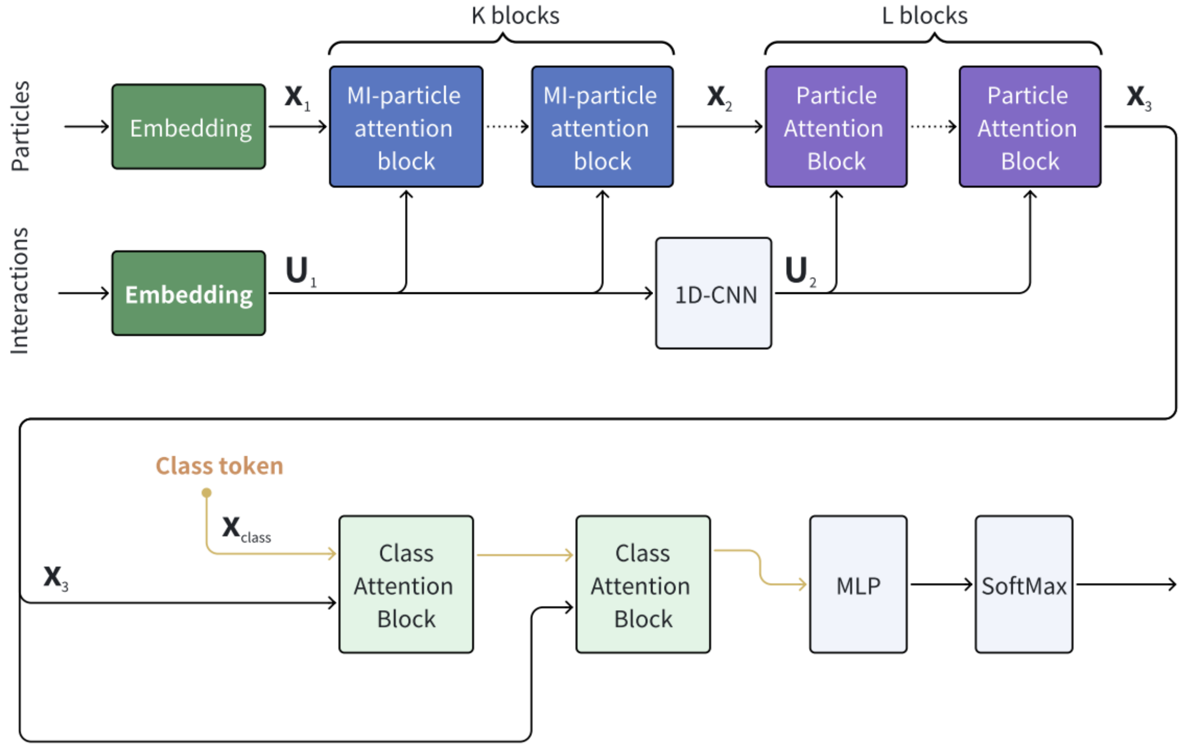


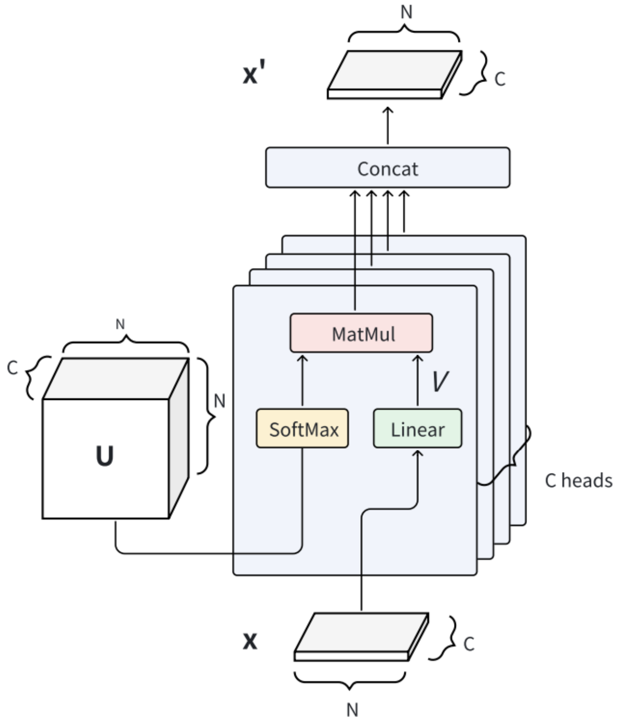


MI-ParT

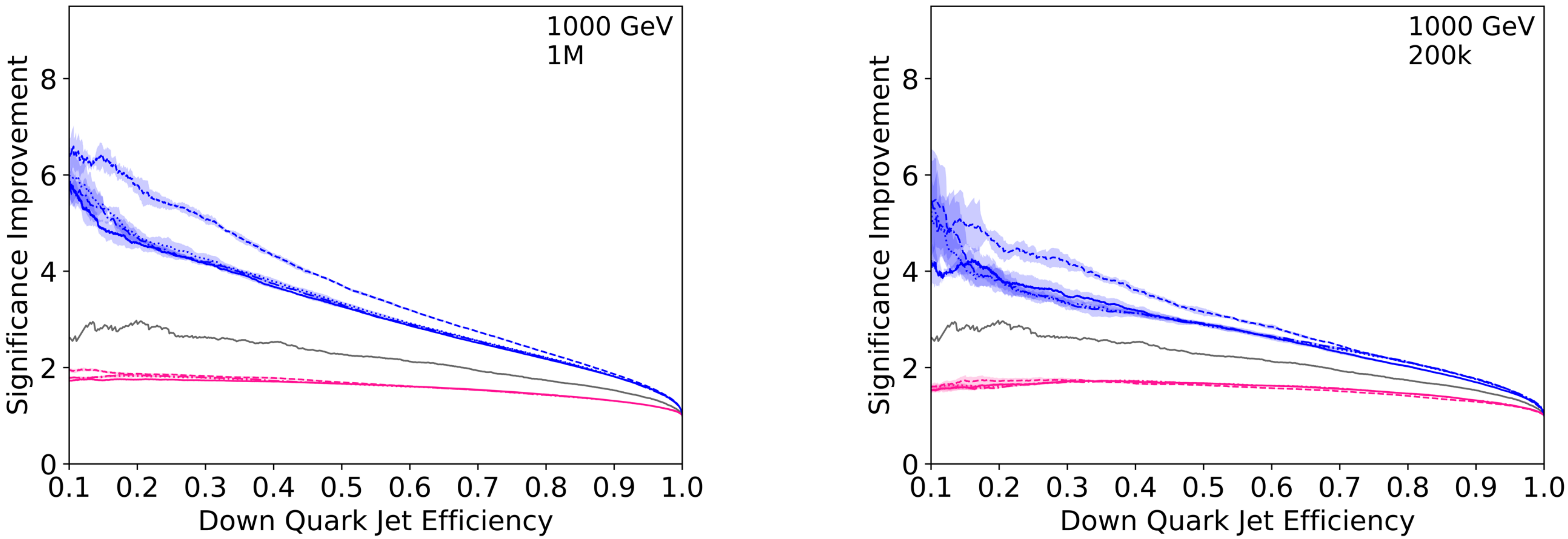
Wu et al., 2407.08682

- Enhances ParT by the modifying attention mechanism to increase the feature dimensions of the interaction embedding





Performance dependence on dataset size

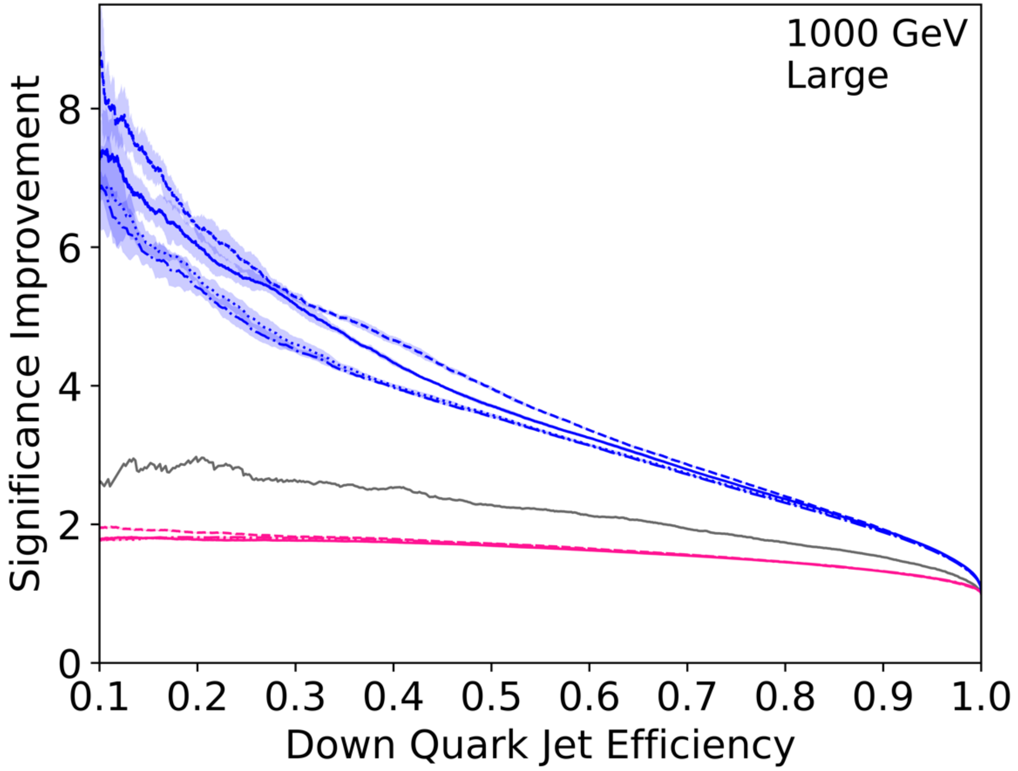


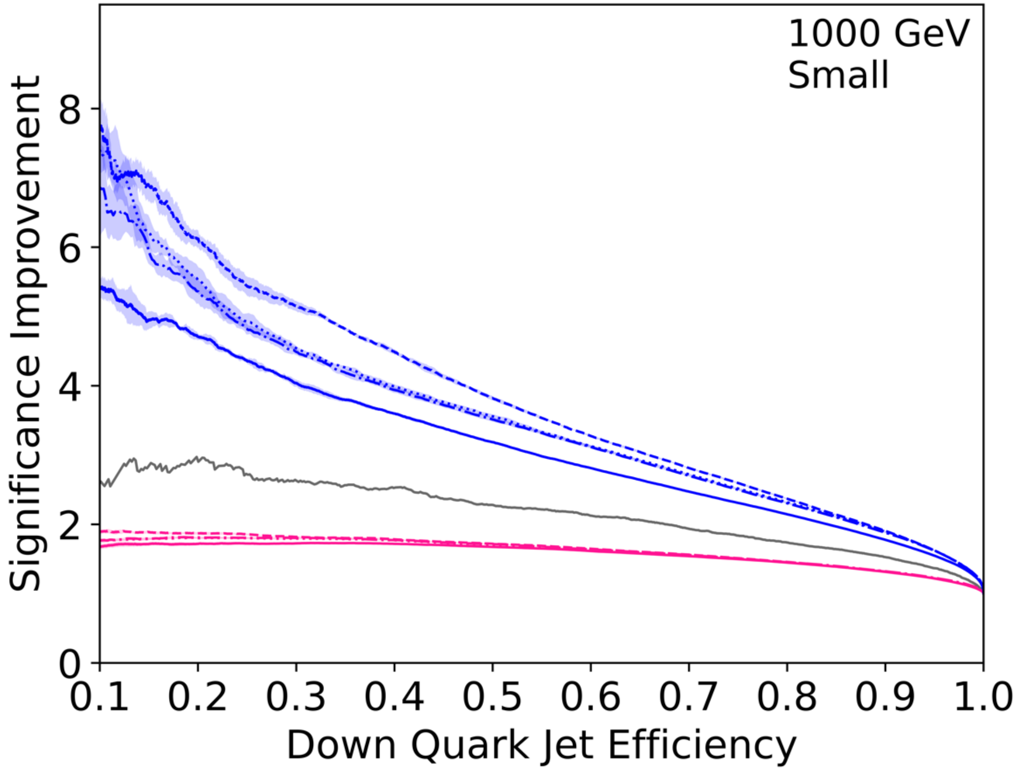
Blue – PID + overall jet charge, κ = 0.2

Pink – overall jet charge, κ = 0.2

Grey – CNN, κ = 0.1 [Fraser and Schwartz, 1803.08066]

Performance dependence on network size





Blue – PID + overall jet charge, κ = 0.2

Pink – overall jet charge, κ = 0.2

Grey – CNN, κ = 0.1 [Fraser and Schwartz, 1803.08066]